Loss of cardio-protective effects at the *ADAMTS7* locus due to gene-smoking interactions

Danish Saleheen PhD,[1,2] Wei Zhao MSc,[1] Robin Young PhD,[3] Christopher P Nelson PhD,[4] WeangKee Ho PhD,[3] Jane Ferguson PhD,[5] Asif Rasheed MBBS,[2] Kristy Ou BS,[5] Sylvia T Nurnberg PhD,[57] Robert C. Bauer PhD,[76] Anuj Goel MSc,[6] Ron Do PhD,[7,8] Alexandre FR Stewart PhD,[9] Jaana Hartiala PhD,[10] Weihua Zhang PhD,[11,12] Gudmar Thorleifsson PhD,[50,51] Rona J Strawbridge PhD,[13] Juha Sinisalo PhD,[61] Stavroula Kanoni PhD,[15] Sanaz Sedaghat PhD,[16] Eirini Marouli PhD,[15,17] Kati Kristiansson PhD,[18] Jing Hua PhD,[19] Robert Scott PhD,[19] Dominique Gauguier PhD,[20] Svati Shah MD,[21] Albert Vernon Smith PhD,[22,36] Natalie R van Zuydam PhD,[23] Amanda J Cox PhD,[24] Christina Willenborg PhD,[25,26], Thorsten Kessler MD,[27,28] Lingyao Zeng PhD,[27,29] Michael A Province PhD,[30] Andrea Ganna PhD,[31,32] Lars Lind PhD,[33] Nancy L. Pedersen PhD,[34] Charles C White PhD,[35] Anni Joensuu MSc,[18,37] Marcus E Kleber PhD,[38] Alistair S Hall PhD,[39] Winfried März PhD,[40] Veikko Salomaa PhD,[18] Christopher O'Donnell MD,[41] Erik Ingelsson PhD,[42,43] Mary F Feitosa PhD,[30] Jeanette Erdmann PhD,[25,26] Donald W. Bowden PhD,[24] Colin N Palmer PhD,[23] Vilmundur Gudnason PhD,[22,36] Ulf De Faire PhD,[44] Pierre Zalloua PhD,[45] Nick Wareham PhD,[20] John R Thompson PhD,[46] Kari Kuulasmaa PhD,[18] George Dedoussis PhD,[17] Markus Perola PhD,[18,37] Abbas Dehghan PhD,[16] John C Chambers PhD,[11,12,47] Jaspal Kooner MD,[11,47,48] Hooman Allayee PhD,[10] Panos Deloukas PhD,[15,49] Ruth McPherson PhD,[9] Kari Stefansson PhD,[50,51] Heribert Schunkert MD,[27,29] Sekar Kathiresan MD,[52-55] Martin Farrall PhD,[6] EPIC-CVD,[3] Philippe Frossard DSC,[2] Daniel J Rader MD,[56,57] Nilesh Samani MD,[4] PROMIS,[2] CARDIoGRAMplusC4D, Muredach P. Reilly MD.[76]

Collaborators:

Stanley L. Hazen MD,[58] W.H. Wilson Tang MD,[58] Perttu P Salo PhD,[18,37] Marja-Liisa Lokki PhD,[60] Markku S Nieminen PhD,[61] Antti-Pekka Sarin MSc,[18,37] Alun Evans MSc,[63] Jean Ferrières MD,[64] Jarmo Virtamo PhD,[18] Frank Kee PhD,[66] David-Alexandre Trégouët PhD,[67] Dominique Arveiler PhD,[68] Philippe Amouyel PhD,[69] Paolo Brambilla PhD,[70] Annette Peters PhD,[71] Melanie Waldenberger PhD,[71,75] Giovanni Veronesi PhD,[72] Giancarlo Cesana PhD,[73] Satu Männistö PhD,[18] Pekka Jousilahti PhD,[18] Antti M Jula PhD,[18] Kennet Harald PhD,[18] Albert Hofman PhD,[16] Oscar H. Franco PhD,[16] Andre G. Uitterlinden PhD.[74]

1. Department of Biostatistics and Epidemiology, University of Pennsylvania, USA
2. Center for Non-Communicable Diseases, Karachi, Pakistan.
3. Department of Public Health and Primary Care, University of Cambridge, UK
4. Department of Cardiovascular Sciences, University of Leicester, Leicester, LE3 9QP, UK
5. Cardiology Division, Department of Medicine, Vanderbilt University, Nashville, USA
6. Division of Cardiovascular Medicine, Radcliffe Department of Medicine & Wellcome Trust Centre for Human Genetics, University of Oxford, UK

7. The Charles Bronfman Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

8. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

9. Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada

10. Institute for Genetic Medicine and Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033

11. Department of Epidemiology and Biostatistics, Imperial College London, London, UK
12. Department of Cardiology, Ealing Hospital NHS Trust, Middlesex, UK
13. Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
14. Department of Medicine, Department of Cardiology, Helsinki University Central Hospital, Helsinki, Finland
15. William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
16. Department of Epidemiology, Erasmus University Medical center, Rotterdam, The Netherlands
17. Department of Dietetics-Nutrition, Harokopio University, 70 El. VenizelouStr, Athens, Greece
18. National Institute for Health and Welfare, Helsinki, Finland
19. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge School of Clinical Medicine, Box 285, Cambridge Biomedical Campus, Cambridge, CB2 0QQ
20. INSERM, UMRS1138, Centre de Recherche des Cordeliers, Paris, France
21. Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, NC 27710
22. Icelandic Heart Association, Kopavogur, Iceland
23. Medical Research Institute, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK
24. Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC, USA
25. Institut für Integrative und Experimentelle Genomik, Universität zu Lübeck, Lübeck, Germany
26. DZHK (German Research Center for Cardiovascular Research) partner site Hamburg–Lübeck–Kiel, Lübeck, Germany
27. Deutsches Herzzentrum München, Technische Universität München, München, Germany
28. Klinikum rechts der Isar, München, Germany
29. DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, München, Germany
30. Department of Genetics, Washington University School of Medicine, St. Louis, MO
31. Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA
32. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA
33. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden
34. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
35. Department of Biostatistics Boston University School of Public Health Framingham Heart Study 801 Massachusetts Avenue, CT3 Boston, MA 02118
36. Faculty of Medicine, University of Iceland, Reykjavik, Iceland
37. University of Helsinki, Institute for Molecular Medicine, Finland (FIMM)
38. Department of Medicine, Mannheim Medical Faculty, Heidelberg University
39. Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, UK
40. Synlab Academy, Synlab Services GmbH, Mannheim, Germany and Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria
41. National Heart, Lung, and Blood Institute and the Framingham Heart Study, National Institutes of Health, Bethesda, MD
42. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
43. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA
44. Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
45. Lebanese American University, School of Medicine, Beirut 1102 2801, Lebanon
46. Department of Health Sciences, University of Leicester, Leicester, UK
47. Imperial College Healthcare NHS Trust, London, UK
48. Cardiovascular Science, National Heart and Lung Institute, Imperial College London, London, UK
49. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia
50. deCODE Genetics, Sturlugata 8, IS-101 Reykjavik, Iceland
51. University of Iceland, School of Medicine, 101 Reykjavik, Iceland

52. Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA
53. Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA
54 Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA
55. Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA
56. Department of Genetics, University of Pennsylvania, USA
57. Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, USA
58. Departments of Cardiovascular Medicine and Cellular and Molecular Medicine, Cleveland Clinic, Cleveland, OH 44195
60. Transplantation Laboratory, Haartman Institute, Helsinki, Finland
61. Heart and Lung Center HUH, Helsinki University Hospital and Helsinki University, Finland
63. Centre for Public Health, The Queen's University of Belfast, Belfast, Northern Ireland
64. INSERM UMR 1027, Department of Cardiology, Toulouse University School of Medicine, Rangueil Hospital, 31059 Toulouse, France
66. UKCRC Centre of Excellence for Public Health (NI), Queens University of Belfast, Northern Ireland
67. Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1166, F-75013, Paris, France AND Sorbonne Universités, Université Pierre et Marie Curie, Paris, France; and Institute for Cardiometabolism and Nutrition, Paris, France
68. Department of Epidemiology and Public Health, EA3430, University of Strasbourg, Faculty of Medicine, Strasbourg, France
69. Institut Pasteur de Lille, INSERM U744, Université Lille Nord de France, F-59000 Lille, France
70. Laboratory Medicine, Hospital of Desio, Department of Health Sciences, University of Milano, Bicocca, Italy
71. Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany AND German Center for Cardiovascular Disease Research (DZHK e.V.), Munich, Germany Melanie Waldenberger
72. EPIMED Research Center, Department of Clinical and Sperimental Medicine, University of Insubria, Varese, Italy
73. University of Milano, Bicocca, Italy
74. Department of Internal Medicine, Erasmus University Medical center, Rotterdam, The Netherlands
75. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
76. Cardiology Division, Department of Medicine and the Irving Institute for Clinical and Translational Research, Columbia University Medical Center, New York, USA

Corresponding authors:

Danish Saleheen, MBBS, PhD
Assistant Professor of Biostatistics and Epidemiology
University of Pennsylvania
11-134 Translational Research Center
3400 Civic Center Boulevard
Philadelphia, PA 19104
Tel: 215-573-6323
Fax: 215-573-2094
Email: saleheen@mail.med.upenn.edu

Muredach P. Reilly, MBBCH, MSCE
Herbert and Florence Irving Professor of Medicine
Director of the Irving Institute for Clinical and Translational Research
Columbia University
622 West 168th Street, PH10-305
New York, NY 10032
Tel: 212-305-9453
Fax: 212-305-3213
Email: mpr2144@cumc.columbia.edu

ABSTRACT

Background Common diseases such as coronary heart disease (CHD) are complex in etiology. The interaction of genetic susceptibility with lifestyle factors may play a prominent role. However, gene-environment interactions for CHD have been difficult to identify. Here, we investigate interaction of smoking behavior, a potent lifestyle factor, with genotypes that have been shown to associate with CHD risk.

Methods We analyzed data on 60,919 CHD cases and 80,243 controls from 29 studies for gene-smoking interactions for genetic variants at 45 loci previously reported to associate with CHD risk. We also studied 5 loci associated with smoking behavior. Study specific gene-smoking interaction effects were calculated and pooled using fixed-effects meta-analyses. Interaction analyses were declared to be significant at a *P-value* < 1.0×10^{-3} (Bonferroni correction for 50 tests).

Results We identified novel gene-smoking interaction for a variant upstream of the *ADAMTS7* gene. Every T allele of rs7178051 was associated with lower CHD risk by 12% in never-smokers (P-value: 1.3×10^{-16}) compared to 5% in ever-smokers (P-value: 2.5×10^{-4}) translating to a 60% loss of CHD protection conferred by this allelic variation in people who smoked tobacco (*Interaction P-value*: 8.7×10^{-5}). The protective T allele at rs7178051 was also associated with reduced *ADAMTS7* expression in human aortic endothelial cells and lymphoblastoid cell lines. Exposure of human coronary artery smooth muscle cells to cigarette smoke extract led to induction of *ADAMTS7.*

Conclusion Allelic variation at rs7178051 that associates with reduced *ADAMTS7* expression confers stronger CHD protection in “never-smokers” compared to “ever-smokers”. Increased vascular *ADAMTS7* expression may contribute to the loss of CHD protection in smokers.

Key words: Gene-smoking interaction, gene-environment interaction, coronary heart disease, ADAMTS7, smoking.

Word count: 269

Clinical Perspective

1) What is new?

- Using data on 60,919 CHD cases and 80,243 controls, this study conducted gene-environment interaction analyses to investigate effect modification by smoking behavior at established CHD and smoking related loci.
- Cardio-protective effects conferred by allelic variation at the *ADAMTS7* locus attenuated by 60% in people who smoked tobacco compared to those who did not smoke.
- Allelic variation at *ADAMTS7* associated with reduced CHD risk was associated with reduced *ADAMTS7* expression in human aortic endothelial cells and lymphoblastoid cell lines.
- Exposure of human coronary artery smooth muscle cells to cigarette smoke extract led to induction of *ADAMTS7.*

2) What are the clinical implications?

- These human genomic data provide new insights into potential mechanisms of CHD in cigarette smokers.
- Findings from this study also point towards the directional impact of the *ADAMTS7* locus on CHD.
- ADAMTS7 and its substrates have a specific role in cigarette smoking related CHD.
- Inhibition of ADAMTS7 is a novel potential therapeutic strategy for CHD that may have particular benefits in individuals who smoke cigarettes.

INTRODUCTION

Coronary heart disease (CHD) is a complex disorder resulting from the interplay of lifestyle and genetic factors.[1, 2] Yet, gene-environment interactions for CHD have been difficult to identify. Cigarette smoking is one of the strongest lifestyle risk factors for CHD but the underlying molecular mechanisms of CHD in humans who smoke remain uncertain.[3-5] Cigarette smoking accounts for one-fifth of all CHD events globally and is responsible for ~1.6 million deaths attributable to CHD annually.[6] Genome-wide association studies (GWAS) have improved our understanding on the genetic predisposition to both CHD and smoking behavior.[7-10] Joint or interactive effects of genetic variation and smoking behavior in the etiology of CHD, however, remain poorly understood. GWAS can provide new opportunities to investigate gene-smoking interactions.

We hypothesized that genetic predisposition to CHD is modified by cigarette smoking at loci discovered by GWAS to be associated with either CHD or smoking behavior. We conducted a focused experiment at 50 main-effect loci (45 for CHD and 5 for smoking behavior) using genetic data and information on smoking behavior in 60,919 CHD cases and 80,243 controls from 29 different studies. We report novel findings on gene-smoking interactions in CHD. Allelic variation on chr.15q25.1 at *ADAMTS7* is associated with protection from CHD in "never-smokers" with attenuation of this protective effect in people who smoked. Expression studies in relevant vascular cells support a role for ADAMTS7 in smoking induced CHD. These data provide the first insights on the etiology of CHD in cigarette smokers and may present opportunities for targeted therapeutic strategies to lower CHD risk in individuals who smoke cigarettes.

METHODS

Summary of study Design

All studies participating in the CARDIoGRAMplusC4D consortium[7-9] that had information available on smoking status, CHD risk and genotypes at the 50 CHD and smoking behavior-associated loci were invited to participate. The current study had five inter-related components (**eFigure-1)**. First, as part of the quality control, we investigated the association of smoking status with CHD risk within each study. Second, we performed an updated analysis of all the SNPs (± 50 KB) at the 45 established CHD loci to identify the variant with the strongest CHD association in our study population at each established CHD locus. Effect estimates from each study in association with CHD risk were obtained and pooled to identify the strongest CHD associated variant ("lead variant"). Third, we investigated gene-smoking interactions at these 45 CHD loci and at 5 loci related to smoking behavior. Fourth, for loci demonstrating differential CHD associations by smoking status, we mapped the interaction region, examined linkage disequilibrium (LD) across the region and performed conditional analyses to identify independent genetic signals. Finally, for loci exhibiting gene-smoking interaction in CHD, we assessed eQTL data for association of variants with expression of local genes in available datasets and examined expression of these genes in multiple cell types that play prominent roles in smoking-CHD pathobiology.

Harmonization of phenotypes and genotypes

Summary level estimates for each study were shared via a secure FTP site. We used "ever-smoking" as a primary exposure and data were harmonized by uniformly characterizing participants in each study into two categories, "ever-smokers" and "never-smokers". "Ever-smokers" were defined as those who had smoked more than 100 cigarettes in a lifetime. For case-control studies, information on "ever smoking" status collected at the time of enrollment was used for the current analyses; whereas for prospective cohort studies, information on smoking status obtained at the baseline visit was used for the current investigation. CHD was defined based on evidence from angiography or history of verified myocardial infarction (MI), percutaneous coronary interventions (PCI) or coronary artery bypass grafting (CABG) as published in CARDIoGRAMplusC4D projects.[7-9] Genotype data generated through GWAS (directly genotyped or imputed) or cardio-metabochip (directly genotyped only) arrays were obtained from each study and all genetic data were aligned using the build-37 reference panel. Imputed SNPs were removed if they had any of the following: (i) a minor allele frequency of <1%; (ii) info score of <0.90; or (iii) confidence score <0.90. For each study, GWAS data were imputed using the Phase II CEU HapMap reference population.[11] Standard

quality control criteria were applied by each participating study, as described previously.[7] All participating studies in the CARDIoGRAMplusC4D consortium were approved by their locally relevant institutional review boards and all participants gave written informed consent before their enrollment in each study. [7-9]

STATISTICAL ANALYSIS

Gene-smoking interaction analyses

Initial quality control and association of established CHD loci with CHD risk: As part of an initial quality control, effect estimates from each study were obtained for "ever-smoking" status and CHD risk using a case-control logistic regression model adjusted for age and sex. Each participating study also assessed and, if needed, controlled for population stratification by including principal components as covariates in the regression model as described earlier.[7-9] To identify variant(s) with the most significant association with CHD risk at established CHD loci in our study population, logistic regression analyses were conducted by each participating study for all the SNPs flanking (±50 kb) the lead variant previously reported at each CHD locus. Effect estimates and standard errors were obtained and meta-analyzed using a fixed-effects inverse variance approach. All lead variants identified through these analyses were further investigated for gene-smoking interactions in CHD. One lead variant per locus was selected for primary gene-smoking interaction analyses.

Investigation of the APOE locus: Although *APOE* has been recently established as a GWAS locus,[7] previous studies prior to GWAS have suggested that CHD risk is higher among carriers of the ε4 allele at the *APOE* locus in smokers than in non-smokers.[12-14] Because the ε2, ε3 and ε4 alleles at the *APOE* locus are not captured by the GWAS platform, we specifically conducted genotyping for rs429358 and rs7412 variants to capture the three epsilon (ε) alleles in 13,822 participants (including 7,286 first-onset myocardial infarction cases) in the PROMIS study.[15]

Gene-smoking interaction analyses at CHD and smoking loci: To assess gene-smoking interactions, analyses were conducted within each study, adjusted for age, sex and other study specific covariates (e.g., principal components), and variants were analyzed in association with CHD separately in "ever-smokers" and "never-smokers". Results from the two groups were then used to test for interaction within each study. For the 50 variants, an interaction test statistic was calculated within each study using the following equation as adapted from *Teslovich TM* et.al.[16]

$$\frac{(\beta n - \beta e)}{\sqrt{SEn^2 + SEe^2}}$$

where β_n and β_e are the beta coefficients for the SNP in never-smokers and ever-smokers respectively, SEn and SEe are the standard errors for the log-ORs estimated for never-smokers and ever-smokers, respectively. Study specific interaction beta(s) and se(s) were calculated within each study and were pooled across studies using a fixed-effects meta-analysis. Interaction analyses were declared to be significant at a P-value of $<1.0 \times 10^{-3}$ (Bonferroni correction for 50 tests).

Conditional analyses on chr.15q25.1: At chr.15q25.1, we observed two variants exhibiting gene-smoking interactions for CHD. The proximity of these two signals raised the possibility that the observed interactions may represent a single interaction locus across the entire region. To investigate this possibility we undertook conditional analyses using an approximate conditional and joint analyses approach, also known as GCTA (Genome-wide Complex Trait Analysis), as described previously.[17-22] Briefly, this method leverages summary-level statistics from a meta-analysis and uses LD corrections between SNPs estimated from a reference sample. Such an approach has been shown to yield similar results to that obtained from conditional analyses conducted on individual participant data and has been successfully implemented in several other studies that have fine-mapped loci for other complex traits.[17-22] Using this approach, we first conducted separate conditional analyses at the chr.15q25.1 locus to identify main-effect variant(s) independently associated with CHD and smoking behavior, respectively. We used the meta-analyzed data for CHD main effects in the CARDIoGRAMplus4D consortium to identify SNPs independently associated with CHD risk and we used the genetic meta-analysis data from the Tobacco and Genetics Consortium (TGC) in 140,000 participants to identify variants independently associated with smoking behavior. We then estimated the effects of these independent variants on CHD risk stratified by smoking status and mutually adjusted the effects of these variants for each other.

Analysis of eQTLs and regulatory features at the chr15q25.1 gene-smoking interaction locus

eQTL analyses: We mined publicly available databases to identify genotype-related expression differences (eQTLs) in *ADAMTS7* and the *CHRNB4-A3-A5* gene cluster in order to understand the directionality of the association of expression of these genes with CHD and smoking behavior. Specifically, we investigated data available from the GTEx consortium,[23] the HapMap consortium (restricted to European populations), and the Multiple Tissue Human Expression Resource (MuTHER).[24] We also analyzed expression data in 147 donor HAoEC lines.[25] We used a nominal P-value of 0.002 to account for multiple testing involved in the eQTL analyses.

Regulatory features of the chr. 15q25.1 region: Data from ENCODE[26] were explored as described in eMethods. ChIP-seq experiments were performed on confluent HCASMC (Cell Applications 350-05a & Lonza CC-2583; cultured in SmGM-2 BulletKit media; Lonza) as described.[27] TCF21 (Abcam ab49475), Jun (Santa Cruz Biotechnology sc-1694), JunD (Santa Cruz Biotechnology sc-74), and CEBP (Santa Cruz Biotechnology sc-150) transcription factor binding was interrogated and H3K27ac data were acquired using the same ChIP protocol with an anti-H3K27ac antibody (Abcam; ab4729). Reads were aligned to the human genome (GRCh37p13) using STAR.[28]

Analyses of ADAMTS7 and _CHRNB4-A3-A5_ gene expression in vascular cells and tissues

ADAMTS7 and *CHRNB4-A3-A5* gene expression in vascular cells: *ADAMTS7* and *CHRNB4-A3-A5* mRNA levels were measured in cultured human coronary artery smooth muscle cells (HCASMC; Lonza CC-2583, Lonza Walkersville, MD), human coronary artery endothelial cells (HCAEC, Lonza CC-2585), human aortic smooth muscle cells (HAoSMC, Lonza CC-2571), human aortic endothelial cells (HAoEC, Lonza CC-2535), human aortic adventitial fibroblasts (HAoAF, Lonza CC-7014), and human acute monocytic leukemia cell line (THP-1, ATCC TIB-202). Further details are provided in eMethods.

ADAMTS7 and *CHRNB4-A3-A5 gene* expression in response to cigarette smoke extract: HCASMC were grown to confluence and cigarette smoke extract experiments performed at passage-7. Cigarette smoke extract was custom-prepared by Arista Laboratories (Richmond, VA). Briefly, the condensate was generated by smoking Marlboro Red King Size Hard Pack cigarettes on an analytical smoke machine under International Organization for Standardization smoking conditions. The smoke condensate was collected on 92 mm filter pads and extracted from each pad in DMSO by shaking to obtain a solution of ~20 mg/mL final concentration of the total particulate matter. Serum starved (24 hrs) HCASMC were treated with 0.5% or 1.0% cigarette smoke extract (v/v) for 4, 12, and 24 hrs in serum reduced conditions (0.5% FBS in DMEM). Details on RNA preparation and q-PCR are provided in eMethods.

RESULTS

Description of the participating studies

Of the 37 studies participating in the CARDIoGRAMplusC4D consortium, information on "ever-smoking" was available in 30 studies, yielding a total sample size of 60,919 CHD cases and 80,243 controls. All studies recruited participants of European ancestry, except PROMIS (South Asian),[15] LOLIPOP (South Asian)[29] and FGENTCARD (Lebanese).[30] Number of CHD cases and controls and percentages that were "ever-smokers" are provided in **eTable 1**. As expected, in all the participating studies, association of "ever-smoking" status with CHD risk was directionally consistent with an increased risk of CHD **(eFigure 2)**.

New variants associated with CHD at established loci

eFigure 3 and **eTable 2** present effect estimates for the association with CHD for (i) the most significant variant that we identified at known CHD loci in the current CARDIoGRAMplusC4D consortium analysis as well as for (ii) the top SNP previously reported at each of these established CHD loci. Of the 45 established CHD loci, we identified 32 for which we discovered a more statistically significant SNP in association with CHD risk in our dataset than the prior reported top variant. All of these 32 SNPs were in moderate to high LD (r^2 >0.6) with the previously published variants.[7-9] In our primary gene-smoking interaction analyses, at each of the CHD loci, we, therefore, used the SNP with the most significant CHD association **(eFigure 3** and **eTable 2)**. Because the smoking behavior phenotype (captured as cigarettes per day [CPD]) was not available in all CARDIoGRAMplusC4D studies, we used the top variant previously reported for CPD[10] at each locus (**eFigure 4**).

Analyses of the APOE locus.

The effect of rs6857, the lead CHD variant at the *APOE* locus, was similar in "ever-smokers" compared to "never-smokers" (**eTable 3)**. Specifically, the CHD OR for the T allele at rs6857 was found to be 1.10 (P-value 7.93×10^{-4}) in "never-smokers" (12,159 CHD cases and 22,932 controls) which was quantitatively similar to the CHD OR of 1.09 (P-value: 8.68×10^{-5}) observed in "ever-smokers" (23,753 CHD cases and 24,019 controls) (interaction P-value: 0.76) (**eFigure 5a**). Investigation in the PROMIS study of the *APOE* epsilon genotypes yielded consistent findings; the OR for CHD among ε4 carriers in "never-smokers" was 1.13 compared to the CHD OR of 1.07 observed in "ever-smokers" (interaction P-value: 0.82) (**eFigure 5a**).

Novel gene-smoking interaction effects on CHD at chromosome 15q25.1

Of the 50 loci, we identified effect-modification by "ever-smoking" status on CHD risk for the lead variants at two distinct loci, rs7178051, in proximity of *ADAMTS7* (an established CHD locus), and rs1051730, in proximity of *CHRNB4-A3-A5* (an established smoking behavior locus) (**eTable 3**). Although associated with different traits and located in distinct LD blocks, these two variants reside ~224 KBs apart on chr.15q25.1 and are in weak linkage disequilibrium (LD) (r^2 = 0.22), raising the question of whether these two chr.15q25.1 gene-smoking interactions on CHD are independent of each other.

At the *ADAMTS7* CHD locus, the T allele at the rs7178051 variant was found to be more strongly and inversely associated with CHD risk in never-smokers (OR: 0.88; P-value: 7.02×10^{-16}) compared to a much weaker effect in ever-smokers (OR: 0.95; P-value: 8.64×10^{-4}) (P-value of interaction: 8.57×10^{-5}) (**Table 1**). Thus, the protective impact of the rs7178051 T allele observed in never-smokers was halved in people who smoked (**Figure-1**). This difference is not related to power differences within strata because for this variant, there were less data available in the never-smoking group (21,232 CHD cases and 38,713 controls) compared to the ever-smoking group (39,585 CHD cases and 40,749 controls). There was no substantial evidence of heterogeneity for the interaction beta across the participating studies (Heterogeneity chi-squared = 36.23 (d.f. = 25); P-value for the χ^2 test of heterogeneity = 0.06; I^2 = 31.0%; tau-squared (τ^2 = 0). We further conducted sensitivity analyses using a random effect model; the results remained unchanged and the interaction beta remained signficiant (**eFigure 5b**). Although the frequency of rs7178051 was 39% in Europeans compared to 28% in South Asians, further analyses stratified by ancestry (i.e., European versus non-Europeans) showed similar results (**eFigure 5c**). Other variants discovered through prior CHD GWAS at the *ADAMTS7* locus (e.g., rs7173743, rs4380028, rs3825807) were in moderate to high LD (r^2 >0.50) with rs7178051 and were also found to display a similar pattern of gene-smoking interaction effects (**Table 1**).

At the *CHRNB4-A3-A5* smoking locus, the A allele at the rs1051730 variant had an inverse trend (not significant after adjustment) of association with CHD in never-smokers (OR: 0.96; P-value: 1.56×10^{-2}) and a positive trend (not significant after adjustment) in ever-smokers (OR: 1.03; P-value: 1.53×10^{-2}) (P-value of interaction: 2.37×10^{-4}) (**Table 1 and eTable 3**). For this variant, data on 20,559 CHD cases and 38,198 controls were available in the never-smoking group whereas 38,923 CHD cases and 40,371 controls were available in the ever-smoking group. Similar gene-smoking

interaction patterns were observed for other variants (e.g., rs2036527, rs8034191) that have been previously reported for CPD behavior at the *CHRNB4-A3-A5* gene cluster (**Table 1**).

Further interrogation of the chr15q21.1 region encompassing rs7178051 and rs1051730 across three distinct LD blocks **(Figure 1)** revealed multiple additional variants for which we observed gene-smoking interactions in CHD (**Table 1** and **Figure 1**). Indeed, several SNPs (e.g., rs7178051, rs10083696, rs7176187, rs6495335, rs4887077) had genome-wide significant associations with CHD in "never-smokers" but had weaker and less significant associations with CHD in "ever-smokers" (**Figure 1**). Alleles clustered specifically around *ADAMTS7* rather than at the *CHRNB4-A3-A5* genes appear to be protective of CHD in "never-smokers" but have attenuated protective effects in "ever-smokers" (**Figure 2**).

Conditional analyses

To investigate the possibility that the two chr.15q25.1 gene-smoking interactions might represent a single interaction locus across the entire region we undertook an approximate conditional and joint analyses[17-22] using summary data derived from CARDIoGRAMplus4D for CHD and from the TGC for smoking behavior. In-addition to rs7178051, we identified one other variant, rs11072794 in low LD with rs7178051 (r^2=0.20) that was associated independently with CHD **(Figure 3a;** red triangles**)** (**Figure 3b & eFigure 6b;** red triangles). We also confirmed two variants (rs1051730 and rs684513) located in two different LD blocks that were independently associated with smoking behavior in the TGC data[10] **(Figure 3d & eFigure 6b;** grey circles**)**.

In analyses of the CHD variants, both rs7178051 and rs11072794 remained strongly associated with CHD after adjusting for the top CPD variants (rs1051730 and rs684513) (**Figure 3d**, red triangles) whereas their weak association with CPD was lost after adjusting for the top CPD variants (**Figure 3d;** grey circles); e.g., the P-value for rs7178051 association with CPD was $1x10^{-5}$ in unadjusted analyses but attenuated to 0.55 after adjusting for rs1051730 and rs684513. In analyses of the CPD variants, both rs1051730 and rs684513 remained strongly associated with CPD after adjusting for the top CHD variants (rs7178051 and rs11072794) (**Figure 3b,** grey circles) whereas their weak association with CHD was lost after adjusting for the top CHD variants **(Figure 3b,** red triangles**)**. As expected, conditional analyses that included all four of these variants resulted in a null association of the region with both CHD and CPD **(eFigure 6b)**. To underscore the validity of the conditional approach using summary data, we used individual participant data from an expanded PROMIS sample involving 9,025 MI cases and 8,506 controls. We found that the OR conferred by allelic variation at rs7178051 remained associated with MI risk independent of the two

CPD variants (rs1051730 and rs684513) and rs11072794 (the second CHD SNP) (**eFigure 6c**). Conversely, the apparent effect of allelic variation at rs1051730 (the top CPD variant) on CHD risk was lost when we adjusted for the other three variants, rs7178051, rs11072794 and rs684513 (**eFigure 6c**).

Next, using summary level data we examined the association of each of these four variants with CHD risk separately in "ever-smokers" and "never-smokers" while mutually adjusting for the other three variants (**Figure 4 & eFigure 7**). In these analyses, only allelic variation at rs7178051 was found to have independent genome-wide significant effects on CHD in never-smokers. rs7178051 was also the only one of these four variants with significant differences in the effect estimate for gene-CHD associations between the two smoking groups (P-value for the χ^2 test of heterogeneity: 5.4×10^{-5}) after adjusting for the effects of other variants (rs11072794, rs1051730 and rs684513). These conditional analyses suggest that (a) variants located near the *ADAMTS7* gene but not *CHRNB4-A3-A5* genes have independent effects on CHD, (b) a single independent gene-smoking interaction signal for CHD exists on chr.15q.25.1 which is localized at the *ADAMTS7* CHD locus (marked by rs7178051) and (c) an apparent interaction signal observed at the nearby *CHRNB4-A3-A5* CPD locus (marked by rs1051730) is not independent of the *ADAMTS7* (rs7178051) interaction signal.

To assess the robustness of conditional analyses methodology that uses summary level data (i.e., GCTA)[17-22], we conducted sensitivity analyses in the PROMIS dataset (9,025 MI cases and 8,506 controls). We assessed the association of rs7178051 (top CHD SNP) and rs1051730 (top CPD SNP) after mutually adjusting for each other by conducting (i) standard logistic regression using individual participant data and (ii) summary level data in PROMIS using the GCTA method (**eTable 4**). The top CHD SNP was found associated with CHD risk in PROMIS independent of the top CPD variant using both the methods, in-contrast the effect on CHD of the top CPD SNP attenuated sharply when adjusted for the top CHD SNP – the effect estimates obtained using the two methods were very similar (**eTable 4**).

Finally, to further demonstrate that the gene-smoking interaction effect in CHD at rs7178051 is independent of the *CHRNB4-A3-A5* CPD locus, we conducted sensitivity analyses in the PROMIS study by restricting our gene-environment interaction analysis to subjects who do not carry the minor alleles of rs1051730 and rs684513 (the two SNPs associated with CPD) at the *CHRNB4-A3-A5* locus. The T allele at the rs7178051 variant was associated with CHD only in never-smokers (OR: 0.88; P-value: 0.01) compared to a weaker and non-significant association in ever-smokers (OR:

0.94; P-value: 0.21) (**eTable 5**). The effect estimates obtained in each of the categories defined by smoking status in PROMIS were similar to the effect estimates obtained in our overall meta-analyses that utilized data in all participants (**eTable 5**).

Analysis of eQTLs and regulatory features at the chr15q25.1 gene-smoking interaction locus.

We mined publicly available eQTL data from the HapMap consortium,[11] GTEx consortium[23] and the MuTHER consortium[24] as well as data from 147 HAoEC lines[25] to examine the association between mRNA expression of *ADAMTS7* and *CHRN* genes with CHD, CPD and gene-smoking interaction SNPs at the chr15q25.1 locus. SNP-mRNA associations with p-values <0.002 (correction for 20 tests) are presented **(Figure 5)**. The top two CHD variants (rs7178051, rs11072794) are associated with reduced *ADAMTS7* expression (e.g., rs11072794 p=6.01x10^{-21} in MuTHER LCL, n=850; and rs7178051 p=0.0029 in HAoEC, n=147) but have no association with expression of *CHRN* genes in any cell or tissue examined. In contrast, the top two CPD variants (rs1051730 and rs684513) were associated with *CHRN* gene expression (e.g., rs1051730 p=6.9x10^{-7} for CHRNA5 in GTEx skeletal muscle and nerve tissue) but have no association with *ADAMTS7* in these cells or tissues. These findings complement conditional analyses suggesting that gene-CHD and gene-smoking interaction effects on CHD are likely mediated by *ADAMTS7* whereas the smoking behavior effect appears to be mediated through the *CHRNA3-5* gene cluster.

In analysis of data from the ENCODE project[26] and for human aortic tissue in NIH Roadmap Epigenomics project, *ADAMTS7* was associated with RNAseq reads and an active transcription mark, H3K36me3, whereas *CHRN* genes had low/absent RNAseq reads and were positive for repressive marks, H3K27me3 and H3K9me3 **(eFigure 8)**. In HCASMC ChIPseq data, rs7178051 the top CHD and gene-smoking CHD interacting SNP, is located in a region with active regulatory marks H3K4me1 and H3K4me3 as well as transcription factor binding site for TCF21, an important HCASMC transcription factor also associated with CAD. This ChIPseq pattern was observed also in human aortic tissue (**Figure 6)**. These regulatory data suggest active transcription of *ADAMTS7*, but not *CHRN* genes, in vascular cells and aortic tissue and reveal that rs7178051, the lead gene-smoking CHD interaction SNP, overlaps active transcription marks and transcription factor binding regions in HCASMC.

ADAMTS7 and *CHRNB4-A3-A5* expression in vascular cells and their response to cigarette smoke extract

In order to explore which genes at the chr15q25.1 locus are expressed in CHD-relevant vascular cells, we performed q-PCR of *ADAMTS7* and the *CHRNB4-A3-A5* genes in primary human vascular cells and in the THP1 human monocyte cell line **(eFigure 9 & Figure 5)**. Whilst *ADAMTS7* mRNA was expressed abundantly in all vascular cell types, mRNA was below detection or expressed at a very low level for each of the genes in the *CHRNB4-A3-A5* cluster in any of these cell types **(eFigure 9)**. Next, we explored the effect of cigarette smoke extract on gene expression in HCASMC, a cell type of particular relevance to vascular responses to cigarette smoke products[31, 32] as well as to *ADAMTS7* vascular functions in atherosclerosis and CHD.[33] In primary HCASMC, cigarette smoke extract exposure increased *ADAMTS7* mRNA levels by over 2-fold **(Figure 5)** but did not affect expression of the *CHRN* genes (not shown). Thus, in contrast to *CHRN* genes, *ADAMTS7* is both expressed and modulated by cigarette smoke extract in CHD-relevant vascular cells providing biological support for *ADAMTS7,* but not CHRN genes, in the gene-smoking interaction at chr15q25.1.

DISCUSSION

We conducted a gene-environment interaction study at fifty loci associated with either CHD or smoking behavior and found evidence of effect-modification of genotype-related CHD risk by smoking-behavior at the chr.15q21.1 CHD locus. Specifically, we observed highly significant attenuation of the cardio-protective effects conferred by alleles at this locus in people who smoked cigarettes. Conditional analyses identified an LD block located at the *ADAMTS7* gene that accounted for both the main effect on CHD as well as the gene-smoking interactions in CHD. Data from expression and cell studies support our genetic analysis, suggesting that the underlying mechanism relates to genotype differences in the effect of smoking on expression of *ADAMTS7* in vascular tissue.

Our findings have novel mechanistic and clinical implications. These human genomic data provide new insights into the mechanism of CHD in cigarette smokers. Identification of gene-smoking interaction at the chr15q21.1 locus suggests a specific role in smoking-related CHD for ADAMTS7 and its substrates, vascular matrix and vascular smooth muscle cell biology more broadly. Such insights can help to prioritize translational strategies for smoking-related CHD and present opportunities to study lifestyle interventions and pharmacological strategies to lower CHD in individuals who smoke cigarettes. Thus, inhibition of ADAMTS7 represents a novel potential therapeutic strategy for CHD that may have particular benefits in individuals who smoke cigarettes. All smokers should receive counseling for smoking cessation yet such broad public health strategies have failed to reach or impact smoking behavior in a large portion of nicotine-addicted individuals. Our data provides a human genomic context for consideration of targeting specific genetically at-risk individuals via intensified preventive strategies and development of novel pharmacological treatments.

Our study also represents a realistic strategy for performing gene-environment interaction studies using contemporary genetic data. We illustrate that identifying joint effects of genetic and lifestyle factors in CHD requires very large sample sizes, yet such analyses are biologically informative when studies are adequately powered. In this context, an important observation in our large sample is the lack of effect modification by smoking behavior on CHD at the *APOE* locus, a previously reported smoking interaction locus.[12-14] This finding is consistent with a recent meta-analysis that found no evidence of effect modification by smoking for *APOE* genotypes and CHD risk.[34] These studies raise concerns that most prior gene-environment interactions studies in CHD have been prone to the same biases (i.e., limited statistical power and false positive associations) as

candidate gene studies investigating main effects in the pre-GWAS era. The present study differs from previous studies by being much larger and, importantly, it includes genomic and functional follow-up data supporting the plausibility of the observed gene-environment interaction.

ADAMTS7 (or the A disintegrin and metalloproteinase with thrombospondin motifs-7) is a member of the ADAMTS family of secreted zinc metalloproteases.[35, 36] We previously discovered and replicated genetic variation at the *ADAMTS7* locus in association with coronary atherosclerosis and MI.[7-9] Both *in vivo* and *in vitro* studies suggest that ADAMTS7 modulates VSMC phenotype switching and migration and that this may be mediated via cartilage oligomeric matrix protein (COMP) or thrombospondin-1 (TSP-1),[32,33] i.e. putative ADAMTS7 substrates expressed in vascular tissue. Genetic variation at *ADAMTS7,* however, has no relationship with traditional risk factors or mechanistic biomarkers; hence the directional impact of *ADAMTS7* expression on CHD risk and the underlying biological mechanisms have been unclear.[32]

Our gene-smoking interaction analyses provide novel insights into the directional impact of the *ADAMTS7* locus on CHD, the underlying mechanisms of CHD in smokers, and how such findings ultimately might translate towards achieving health benefits in society. Our human eQTL interrogations reveal that common alleles that relate to lower CHD risk at the *ADAMTS7* locus are also associated with reduced *ADAMTS7* expression, implying an atherogenic role of the gene. This is supported by our recent *in vivo* experimental studies; *Adamts7* deficiency protected against diet-induced atherosclerosis in both the *Ldlr*$^{-/-}$ and *ApoE*$^{-/-}$ mouse models, reduced neointima formation following arterial injury, and decreased VSMC migration *in vitro*.[33] In our smoking-stratified analyses, we observed CHD protective effect which was attenuated in smokers. Thus, smoking exposure may overcome the genetic effect of protective alleles that act by reducing *ADAMTS7* expression. Such a possibility is supported by our HCASMC data that reveals increased *ADAMTS7* expression in HCASMC exposed to cigarette smoke extract. These human genome-smoking studies are the first to implicate a specific locus as causal in mediating increased risk of CHD in smokers. Additional translational studies are needed to establish the precise mechanisms of atheroprotection for alleles at the *ADAMTS7* locus, how cigarette smoking impacts these genetic effects, and whether deletion or inhibition of ADAMTS7 *in vivo* attenuates the specific acceleration of atherosclerosis conferred by cigarette smoking.

Strengths and limitations of our study merit consideration. This is a large study that conducted gene-smoking interaction analyses in CHD by using GWAS data. We observed substantial heterogeneity across study samples in our initial quality control analyses of "ever-

smoking" status with CHD risk. This is similar, however, to the heterogeneity reported in a recent meta-analysis that pooled risk ratios from all the past prospective studies with information on association of "ever-smoking" with incident CHD events.[5] We recognize that other smoking related phenotypes are important e.g., "current smoking" may have a more pronounced role than "ever-smoking" in plaque rupture and thrombosis in patients with MI. "Current smoking" status and MI phenotypes were available only in a subset of our studies limiting statistical power. Given the use of multiple studies and meta-analyses of data, we used only one analytical approach to investigate gene-smoking interactions. This approach, however, was feasible and powerful in this large-scale consortium setting. While we used a fixed effects approach in our meta-analyses, a random effects meta-analysis yielded qualitatively similar results (data not shown). The lack or replication is partially offset by a large sample size, consistency across study cohorts and racial groups and supplemental genomic and experimental evidence supporting biological plausibility. This approach is also consistent with recent recommendations[37] which favor use of a powerful discovery experiment using all data rather than reducing power by splitting available dataset for discovery and validation. While our *in vitro* studies support a role for ADAMTS7 in the gene-smoking interaction, it will be important to confirm that *Adamts7* deficiency protect against cigarette-smoke acceleration of atherosclerosis in rodent models.

Our interaction analyses, conditional analyses, eQTL interrogations and cell studies suggest that *ADAMTS7*, but not the *CHRNB4-A3-A5* gene cluster, is likely causal at 15q21.1 for gene-smoking interaction effects in CHD. Yet, analyses are not definitive. Although top interacting SNPs and CHD SNPs (e.g., rs7178051) were associated with *ADAMTS7,* but not *CHRNB4-A3-A5,* expression in LCLs, large-scale eQTL data and allele specific expression data (e.g., via RNA sequencing) are not available for vascular tissues limiting causal inference. In our small HCAEC datasets, we did however find that alleles at rs7178051 associate with *ADAMTS7* expression but not with any *CHRNB4-A3-A5* genes suggesting, at least in one vascular cell type, that the gene-smoking interaction is mediated via *ADAMTS7.*

Conclusions

We provide novel evidence for allelic variation exhibiting gene-smoking interaction in CHD at the chr.15q21.1 locus. The protective effect conferred by variation at this locus in never-smokers is markedly attenuated in people who are ever-smokers. Stepwise conditional analyses, gene expression data in vascular cells, eQTL interrogation, and cigarette smoke extract exposure in HCASMC suggest that *ADAMTS7* accounts for both the gene-smoking interaction in CHD and the CHD main effect on chr.15q21.1. Our findings reveal interactions of genetic variants and key lifestyle determinants in the etiology of CHD, provide new insights into the potential mechanisms of CHD in cigarette smokers, and facilitate precision medicine advances in cigarette-smoking related CHD. Our work motivates future large-scale studies investigating joint effects of genes and environment in CHD using existing complex-disease consortia datasets and genome-wide discovery approaches. This will provide opportunities to detect additional and novel loci displaying gene-environment interactions revealing genetic contexts for targeting intensive lifestyle interventions and novel therapeutics.

Sources of Funding

Dr. Saleheen has received funding from the National Institutes of Health, the Fogarty International, the Wellcome Trust, the British Heart Foundation, Pfizer, Genentech, Regeneron and Eli Lilly pharmaceuticals. This work was supported in part by R01-HL-111694 and K24-HL-107643 from the National Institutes of Health to Dr. Reilly.

PROMIS. Genotyping in PROMIS was funded by the Wellcome Trust, UK and Pfizer. Fieldwork in the PROMIS study has been supported through funds available to investigators at the Center for Non-Communicable Diseases, Pakistan and the University of Cambridge, UK.

EPIC-CVD Consortium. CHD case ascertainment and validation, genotyping, and clinical chemistry assays in EPIC-CVD were principally supported by grants awarded to the University of Cambridge from the EU Framework Programme 7 (HEALTH-F2-2012-279233), the UK Medical Research Council (G0800270) and British Heart Foundation (SP/09/002), the UK National Institute for Health Research Cambridge Biomedical Research Centre, and the European Research Council (268834). Scientists at the EPIC-CVD Coordinating Centre have also been supported by grants from the US National Institutes of Health, Merck, Novartis, GlaxoSmithKline, and Pfizer.

WTCCC. Recruitment of the WTCCC CAD cases was funded by the British Heart Foundation. Collection of controls and genotyping was funded by the Wellcome Trust. Chris Nelson and Nilesh Samani are funded by the British Heart Foundation. Nilesh Samani is a NIHR Senior Investigator.

SCARF-SHEEP. The investigators would like to acknowledge the Swedish Heart-Lung Foundation, the Swedish Research Council, the Strategic Cardiovascular Programme of Karolinska Institutet and the Stockholm County Council, the Strategic support for epidemiological research at Karolinska Institutet and the Stockholm County Council. Rona J Strawbridge is supported by SRP Diabetes Program at Karolinska Institutet.

CARDIOGENICS and THESIAS. Professor Deloukas' work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit which is supported and funded by the National Institute for Health Research. Analysis was supported by BHF grant (Deloukas) RG/14/5/30893**.** Professor Schunkert was supported by grants from the Fondation Leducq (*CADgenomics*: Understanding CAD Genes, 12CVD02), the German Federal

Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (*e:AtheroSysMed*, grant 01ZX1313A-2014), and the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n° HEALTH-F2-2013-601456 (*CVgenes-at-target*). Further grants were received from the DFG as part of the Sonderforschungsbereich CRC 1123 (B2). Thorsten Kessler MD was supported by a DZHK Rotation Grant.

Rotterdam Study. The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam; the Netherlands Organization for Scientific Research; the Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministryof Health, Welfare and Sports; the European Commission; and the Municipality of Rotterdam. Support for genotyping was provided by the Netherlands Organisation of Scientific Research (NOW) Investments (No. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative/Netherlands Consortium for Healthy Aging project No. 050-060-810. Abbas Dehghan is supported by an NWO grant (veni, 916.12.154) and the EUR Fellowship. Oscar H. Franco works in ErasmusAGE, a center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA. Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Expression studies. These studies were supported in part by a Transatlantic Network of Excellence grant 10CVD03 from the Fondation Leducq. Analysis of expression quantitative trait loci (eQTL) in endothelial cells was supported by a Transatlantic Networks of Excellence Award (12CVD02) from Foundation Leducq (to Dr. Jake Lusis and team)

Diabetes Heart Study. This study was supported in part by R01 HL67348, R01 HL092301, and R01 NS058700 to Donald W. Bowden.

Cleveland Clinic Study. This study was supported in part by NIH grants R01ES021801, 3R01ES021801-03S1, and R01ES025786.

Family Heart Study (FamHS). The FamHS is funded by NHLBI grant R01HL117078 grant.

MORGAM. This work has been sustained by the MORGAM Project's funding: European Union FP 7 projects ENGAGE (HEALTH-F4-2007-201413), CHANCES (HEALTH-F3-2010-242244) and BiomarCaRE (278913). This funding has supported central coordination, workshops and part of the

activities of the MORGAM Data Centre, at THL in Helsinki, Finland. MORGAM Participating Centres are funded by regional and national governments, research councils, charities, and other local sources. The PRIME Study was supported by grants from Inserm, Merck Sharp and Dohme-Chibret Laboratory, the French Research Agency and the Foundation Heart and Arteries. We also thank the following organisations that allowed the recruitment of participants for the PRIME: the health screening centres organised by the SocialSecurity of Lille (Institut Pasteur), Strasbourg, Toulouse, and Tourcoing; the occupational medicine services of Haute-Garonne and of the Urban Community of Strasbourg; the Association Inter-entreprises des Services Médicaux du Travail de Lille et environs; the Comité pour le Développement de la Médecine du Travail; the Mutuelle Générale des Postes, Télégraphes et Téléphones du Bas-Rhin; the Laboratoire d'Analyses de l'Institut de Chimie Biologique de la Faculté de Médecine de Strasbourg. We also gratefully acknowledge the teams of the Lille, Strasbourg and Toulouse centres for their dedicate work and relentness energy in following up their cohorts; the contribution of the members of the event validation committees : L Guize; C Morrison; M-T Guillanneuf; and M Giroud and the Alliance Partnership Programme for its financial support. The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. KK was supported by the Orion-Farmos Research Foundation and Academy of Finland (grant number 250207).

DILGOM. This work was enabled through a grant #139635 from the Academy of Finland and a grant from the Finnish Foundation for Cardiovascular Research. The DILGOM project is also supported by the Academy of Finland (grant numbers 136895 and 263836). We are grateful for the THL DNA laboratory for its skillful work to produce the DNA samples used in this study. We also acknowledge the Academy of Finland (136895 and 263836), Funding from Academy of Finland, grant number: 118065, and Juho Vainio Foundation.

Acknowledgements

We would like to thank the CARDIoGRAMplusC4D consortium, the EPIC-CVD and the PROMIS study for contributing data.

PROMIS. We also acknowledge the contributions made by the following: Zeeshan Ozair, Usman Ahmed, Abdul Hakeem, Hamza Khalid, Naeem Khan, Sadiq Khan, Ayaz Ali, Madad Ali, Saeed Ahmed, Muhammad Waqar Khan, Muhammad Razaq Khan, Abdul Ghafoor, Mir Alam, Riazuddin, Muhammad Irshad Javed, Abdul Ghaffar, Tanveer Baig Mirza, Muhammad Shahid, Jabir Furqan, Muhammad Iqbal Abbasi, Tanveer Abbas, Rana Zulfiqar, Muhammad Wajid, Irfan Ali, Muhammad Ikhlaq, Danish Sheikh, Muhammad Imran, Nadeem Sarwar, Adam Butterworth, Matthew Walker and Hannah Lombardi, Shahid Abbas, Faisal Majeed, Saba Akhtar, Nadeem Qamar, Khan Shah Zaman, Zia Yaqoob, Tahir Saghir, Syed Nadeem Hasan Rizvi, Anis Memon, Nadeem Hayyat Mallick, Mohammad Ishaq, Syed Zahed Rasheed, Fazal-ur-Rehman Memon, Khalid Mahmood, and Naveeduddin Ahmed.

EPIC-CVD. We thank all EPIC participants and staff for their contribution to the study, the laboratory teams at the Medical Research Council Epidemiology Unit for sample management and Cambridge Genomic Services for genotyping, Sarah Spackman for data management, and the team at the EPIC-CVD Coordinating Centre for study coordination and administration.

Disclosures

Dr. Saleheen has received funding from Pfizer, Genentech, Regeneron and Eli Lilly pharmaceuticals.

References

(1) Mangino M, Spector T. Understanding coronary artery disease using twin studies. Heart. 2013;99:373-5.

(2) Lusis AJ, Mar R, Pajukanta P. Genetics of atherosclerosis. Annu Rev Genomics Hum Genet. 2004;5:189-218.

(3) Teo KK, Ounpuu S, Hawken S, Pandey MR, Valentin V, Hunt D, Diaz R, Rashed W, Freeman R, Jiang L, Zhang X, Yusuf S; INTERHEART Study Investigators. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. Lancet. 2006;368:647-58.

(4) Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ. 2004;328:1519.

(5) Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. Lancet. 2011;378:1297-305.

(6) Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. Lancet. 2003;362:847-52.

(7) CARDIoGRAMplusC4D Consortium., Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, König IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R; DIAGRAM Consortium.; CARDIOGENICS Consortium., Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M; MuTHER Consortium., Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, van der Schoot CE, Wagner PJ; Wellcome Trust Case Control Consortium., Wells GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jang Y, Kähönen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee JY, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Trégouët DA, Virtamo J, Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen T, Syvänen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J, Palmer CN, Roberts R, Watkins H, Schunkert H, Samani NJ. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013;45:25-33.

(8) Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet. 2011;43:339-44.

(9) Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I; Cardiogenics., Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J; CARDIoGRAM Consortium., Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333-8.

(10) Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet. 2010;42:441-7.

(11) International HapMap 3 Consortium., Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE, Altshuler DM, Gibbs RA, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Dermitzakis E, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Bonnen PE, Gibbs RA, Gonzaga-Jauregui C, Keinan A, Price AL, Yu F, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Schaffner SF, Zhang Q, Ghori MJ, McGinnis R, McLaren W, Pollack S, Price AL, Schaffner SF, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE. Integrating common and rare genetic variation in diverse human populations. Nature. 2010;467:52-8.

(12) Humphries SE, Talmud PJ, Hawe E, Bolla M, Day IN, Miller GJ. Apolipoprotein E4 and coronary heart disease in middle-aged men who smoke: a prospective study. Lancet. 2001;358:115-9.

(13) Grammer TB, Hoffmann MM, Scharnagl H, Kleber ME, Silbernagel G, Pilz S, Tomaschitz A, Lerchbaum E, Siekmeier R, März W. Smoking, apolipoprotein E genotypes, and mortality (the Ludwigshafen RIsk and Cardiovascular Health study). Eur Heart J. 2013;34:1298-305.

(14) Gustavsson J, Mehlig K, Leander K, Strandhagen E, Björck L, Thelle DS, Lissner L, Blennow K, Zetterberg H, Nyberg F. Interaction of apolipoprotein E genotype with smoking and physical inactivity on coronary heart disease risk in men and women. Atherosclerosis. 2012;220:486-92.

(15) Saleheen D, Zaidi M, Rasheed A, Ahmad U, Hakeem A, Murtaza M, Kayani W, Faruqui A, Kundi A, Zaman KS, Yaqoob Z, Cheema LA, Samad A, Rasheed SZ, Mallick NH, Azhar M, Jooma R, Gardezi AR, Memon N, Ghaffar A, Fazal-ur-Rehman, Khan N, Shah N, Ali Shah A, Samuel M, Hanif F, Yameen M, Naz S, Sultana A, Nazir A, Raza S, Shazad M, Nasim S, Javed MA, Ali SS, Jafree M, Nisar MI, Daood MS, Hussain A, Sarwar N, Kamal A, Deloukas P, Ishaq M, Frossard P, Danesh J. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. Eur J Epidemiol. 2009;24:329-38.

(16) Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee JY, Park T, Kim K, Sim X, Twee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X, Luan J, Lamina C, Ziegler A, Zhang W, Zee RY, Wright AF, Witteman JC, Wilson JF, Willemsen G, Wichmann HE, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF, Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM, Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJ, Scuteri A, Scott J, Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruokonen A, Rudan I, Rose LM, Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BW, Pedersen NL, Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery GW, Meitinger T, McPherson R, McCarthy MI, McArdle W, Masson D, Martin NG, Marroni F, Mangino M, Magnusson PK, Lucas G, Luben R, Loos RJ, Lokki ML, Lettre G, Langenberg C, Launer LJ, Lakatta EG, Laaksonen R, Kyvik KO, Kronenberg F, König IR, Khaw KT, Kaprio J, Kaplan LM, Johansson A, Jarvelin MR, Janssens AC, Ingelsson E, Igl W, Kees Hovingh G, Hottenga JJ, Hofman A, Hicks AA, Hengstenberg C, Heid IM, Hayward C, Havulinna AS, Hastie ND, Harris TB, Haritunians T, Hall AS, Gyllensten U, Guiducci C, Groop LC, Gonzalez E, Gieger C, Freimer NB, Ferrucci L, Erdmann J, Elliott P, Ejebe KG, Döring A, Dominiczak AF, Demissie S, Deloukas P, de Geus EJ, de Faire U, Crawford G, Collins FS, Chen YD, Caulfield MJ, Campbell H, Burtt NP, Bonnycastle LL, Boomsma DI, Boekholdt SM, Bergman RN, Barroso I, Bandinelli S, Ballantyne CM, Assimes TL, Quertermous T, Altshuler D, Seielstad M, Wong TY, Tai ES, Feranil AB, Kuzawa CW, Adair LS, Taylor HA Jr, Borecki IB, Gabriel SB, Wilson JG, Holm H, Thorsteinsdottir U, Gudnason V, Krauss RM, Mohlke KL, Ordovas JM, Munroe PB, Kooner JS, Tall AR, Hegele RA, Kastelein JJ, Schadt EE, Rotter JI, Boerwinkle E, Strachan DP, Mooser V, Stefansson K, Reilly MP, Samani NJ, Schunkert H, Cupples LA, Sandhu MS, Ridker PM, Rader DJ, van Duijn CM, Peltonen L, Abecasis GR, Boehnke M, Kathiresan S. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707-13..

(17) Yang J, Ferreira T, Morris AP, Medland SE; Genetic Investigation of ANthropometric Traits (GIANT) Consortium.; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium., Madden PA, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy MI, Hirschhorn JN, Goddard ME, Visscher PM. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012;44:369-3.

(18) Yang J, Lee SH, Goddard ME, Visscher PM. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. Methods Mol Biol. 2013; 1019:215-36.

(19) Keller MF, Ferrucci L, Singleton AB. Genome-Wide Analysis of the Heritability of Amyotrophic Lateral Sclerosis. JAMA Neurol. 2014;71:1123-34.

(20) van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, van der Spek RA, Võsa U, de Jong S, Robinson MR, Yang J, Fogh I, van Doormaal PT, Tazelaar GH, Koppers M, Blokhuis AM, Sproviero W, Jones AR, Kenna KP, van Eijk KR, Harschnitz O, Schellevis RD, Brands WJ, Medic J, Menelaou A, Vajda A, Ticozzi N, Lin K, Rogelj B, Vrabec K, Ravnik-Glavač M, Koritnik B, Zidar J, Leonardis L, Grošelj LD, Millecamps S, Salachas F, Meininger V, de Carvalho M, Pinto S, Mora JS, Rojas-García R, Polak M, Chandran S, Colville S, Swingler R, Morrison KE, Shaw PJ, Hardy J, Orrell RW, Pittman A, Sidle K, Fratta P, Malaspina A, Topp S, Petri S, Abdulla S, Drepper C, Sendtner M, Meyer T, Ophoff RA, Staats KA, Wiedau-Pazos M, Lomen-Hoerth C, Van Deerlin VM, Trojanowski JQ, Elman L, McCluskey L, Basak AN, Tunca C, Hamzeiy H, Parman Y, Meitinger T, Lichtner P, Radivojkov-Blagojevic M, Andres CR, Maurel C, Bensimon G, Landwehrmeyer B, Brice A, Payan CA, Saker-Delye S, Dürr A, Wood NW, Tittmann L, Lieb W, Franke A, Rietschel M, Cichon S, Nöthen MM, Amouyel P, Tzourio C, Dartigues JF, Uitterlinden AG, Rivadeneira F, Estrada K, Hofman A, Curtis C, Blauw HM, van der Kooi AJ, de Visser M, Goris A, Weber M, Shaw CE, Smith BN, Pansarasa O, Cereda C, Del Bo R, Comi GP, D'Alfonso S, Bertolin C, Sorarù G, Mazzini L, Pensato V, Gellera C, Tiloca C, Ratti A, Calvo A, Moglia C, Brunetti M, Arcuti S, Capozzo R, Zecca C, Lunetta C, Penco S, Riva N, Padovani A, Filosto M, Muller B, Stuit RJ; PARALS Registry.; SLALOM Group.; SLAP Registry.; FALS Sequencing Consortium.; SLAGEN Consortium.; NNIPPS Study Group., Blair I, Zhang K, McCann EP, Fifita JA, Nicholson GA, Rowe DB, Pamphlett R, Kiernan MC, Grosskreutz J, Witte OW, Ringer T, Prell T, Stubendorff B, Kurth I, Hübner CA, Leigh PN, Casale F, Chio A, Beghi E, Pupillo E, Tortelli R, Logroscino G, Powell J, Ludolph AC, Weishaupt JH, Robberecht W, Van Damme P, Franke L, Pers TH, Brown RH, Glass JD, Landers JE, Hardiman O, Andersen PM, Corcia P, Vourc'h P, Silani V, Wray NR, Visscher PM, de Bakker PI, van Es MA, Pasterkamp RJ, Lewis CM, Breen G, Al-Chalabi A, van den Berg LH, Veldink JH. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet. 2016;48:1043-8.

(21) Torres JM, Gamazon ER, Parra EJ, Below JE, Valladares-Salgado A, Wacher N, Cruz M, Hanis CL, Cox NJ. Cross-tissue and tissue-specific eQTLs: partitioning the heritability of a complex trait. Am J Hum Genet. 2014;95:521-34.

(22) Bailey JN, Loomis SJ, Kang JH, Allingham RR, Gharahkhani P, Khor CC, Burdon KP, Aschard H, Chasman DI, Igo RP Jr, Hysi PG, Glastonbury CA, Ashley-Koch A, Brilliant M, Brown AA, Budenz DL, Buil A, Cheng CY, Choi H, Christen WG, Curhan G, De Vivo I, Fingert JH, Foster PJ, Fuchs C, Gaasterland D, Gaasterland T, Hewitt AW, Hu F, Hunter DJ, Khawaja AP, Lee RK, Li Z, Lichter PR, Mackey DA, McGuffin P, Mitchell P, Moroi SE, Perera SA, Pepper KW, Qi Q, Realini T, Richards JE, Ridker PM, Rimm E, Ritch R, Ritchie M, Schuman JS, Scott WK, Singh K, Sit AJ, Song YE, Tamimi RM, Topouzis F, Viswanathan AC, Verma SS, Vollrath D, Wang JJ, Weisschuh N, Wissinger B, Wollstein G, Wong TY, Yaspan BL, Zack DJ, Zhang K, Study EN; ANZRAG Consortium., Weinreb RN, Pericak-Vance MA, Small K, Hammond CJ, Aung T, Liu Y, Vithana EN, MacGregor S, Craig JE, Kraft P, Howell G, Hauser MA, Pasquale LR, Haines JL, Wiggs JL. Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. Nat Genet. 2016;48:189-94

(23) GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45:580-5.

(24) Grundberg E, Small KS, Hedman ÅK, Nica AC, Buil A, Keildson S, Bell JT, Yang TP, Meduri E, Barrett A, Nisbett J, Sekowska M, Wilk A, Shin SY, Glass D, Travers M, Min JL, Ring S, Ho K, Thorleifsson G, Kong A, Thorsteindottir U, Ainali C, Dimas AS, Hassanali N, Ingle C, Knowles D, Krestyaninova M, Lowe CE, Di Meglio P, Montgomery SB, Parts L, Potter S, Surdulescu G, Tsaprouni L, Tsoka S, Bataille V, Durbin R, Nestle FO, O'Rahilly S, Soranzo N, Lindgren CM, Zondervan KT, Ahmadi KR, Schadt EE, Stefansson K, Smith GD, McCarthy MI, Deloukas P, Dermitzakis ET, Spector TD; Multiple Tissue Human Expression Resource (MuTHER) Consortium..

Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet. 2012;44:1084-9.

(25) Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA, Lusis AJ. Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. J Lipid Res. 2013;54:1894-905.

(26) ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 2011;9:e1001046.

(27) Miller CL, Anderson DR, Kundu RK, Raiesdana A, Nürnberg ST, Diaz R, Cheng K, Leeper NJ, Chen CH, Chang IS, Schadt EE, Hsiung CA, Assimes TL, Quertermous T. Disease-related growth factor and embryonic signaling pathways modulate an enhancer of TCF21 expression at the 6q23.2 coronary heart disease locus. PLoS Genet. 2013;9:e1003652.

(28) Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15-21.

(29) Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, Frossard P, Been LF, Chia KS, Dimas AS, Hassanali N, Jafar T, Jowett JB, Li X, Radha V, Rees SD, Takeuchi F, Young R, Aung T, Basit A, Chidambaram M, Das D, Grundberg E, Hedman AK, Hydrie ZI, Islam M, Khor CC, Kowlessur S, Kristensen MM, Liju S, Lim WY, Matthews DR, Liu J, Morris AP, Nica AC, Pinidiyapathirage JM, Prokopenko I, Rasheed A, Samuel M, Shah N, Shera AS, Small KS, Suo C, Wickremasinghe AR, Wong TY, Yang M, Zhang F; DIAGRAM; MuTHER, Abecasis GR, Barnett AH, Caulfield M, Deloukas P, Frayling TM, Froguel P, Kato N, Katulanda P, Kelly MA, Liang J, Mohan V, Sanghera DK, Scott J, Seielstad M, Zimmet PZ, Elliott P, Teo YY, McCarthy MI, Danesh J, Tai ES, Chambers JC. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet. 2011;43:984-9.

(30) Saade S, Cazier JB, Ghassibe-Sabbagh M, Youhanna S, Badro DA, Kamatani Y, Hager J, Yeretzian JS, El-Khazen G, Haber M, Salloum AK, Douaihy B, Othman R, Shasha N, Kabbani S, Bayeh HE, Chammas E, Farrall M, Gauguier D, Platt DE, Zalloua PA; FGENTCARD consortium. Large scale association analysis identifies three susceptibility loci for coronary artery disease. PLoS One. 2011;6:e29427.

(31) Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, Viñuela A, Grundberg E, Nelson CP, Meduri E, Buil A, Cambien F, Hengstenberg C, Erdmann J, Schunkert H, Goodall AH, Ouwehand WH, Dermitzakis E, Spector TD, Samani NJ, Deloukas P. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. Epigenetics. 2014;9:1382-96.

(32) Starke RM, Ali MS, Jabbour PM, Tjoumakaris SI, Gonzalez F, Hasan DM, Rosenwasser RH, Owens GK, Koch WJ, Dumont AS. Cigarette smoke modulates vascular smooth muscle phenotype: implications for carotid and cerebrovascular disease. PLoS One. 2013;8:e71954.

(33) Bauer RC, Tohyama J, Cui J, Cheng L, Yang J, Zhang X, Ou K, Paschos GK, Zheng XL, Parmacek MS, Rader DJ, Reilly MP. Knockout of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. Circulation. 2015;131:1202-13.

(34) Holmes MV, Frikke-Schmidt R, Melis D, Luben R, Asselbergs FW, Boer JM, Cooper J, Palmen J, Horvat P, Engmann J, Li KW, Onland-Moret NC, Hofker MH, Kumari M, Keating BJ, Hubacek JA,

Adamkova V, Kubinova R, Bobak M, Khaw KT, Nordestgaard BG, Wareham N, Humphries SE, Langenberg C, Tybjaerg-Hansen A, Talmud PJ. A systematic review and meta-analysis of 130,000 individuals shows smoking does not modify the association of APOE genotype on risk of coronary heart disease. Atherosclerosis. 2014;237:5-12.

(35) Pu X, Xiao Q, Kiechl S, Chan K, Ng FL, Gor S, Poston RN, Fang C, Patel A, Senver EC, Shaw-Hawkins S, Willeit J, Liu C, Zhu J, Tucker AT, Xu Q, Caulfield MJ, Ye S. ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. Am J Hum Genet. 2013;92:366-74.

(36) Wang L, Zheng J, Bai X, Liu B, Liu CJ, Xu Q, Zhu Y, Wang N, Kong W, Wang X. ADAMTS-7 mediates vascular smooth muscle cell migration and neointima formation in balloon-injured rat arteries. Circ Res. 2009;104:688-98.

(37) Hutter CM, Mechanic LE, Chatterjee N, Kraft P, Gillanders EM; NCI Gene-Environment Think Tank. Gene-environment interactions in cancer epidemiology: a National Cancer Institute Think Tank report. Genet Epidemiol. 2013;37:643-57.

Figure Legends

Figure 1. (a) Regional association analyses at the chromosome 15q25.1 locus in association with CHD risk stratified by smoking status. Association P-values for genetic variants with CHD risk in "never-smokers" (green squares) and "ever-smokers" (red triangles). (b) Longitudinal bars represent gene-smoking CHD interaction P-values at the chromosome 15q25.1 locus; bars in blue are P-values for variants listed in Table-1 and each variant has been assigned a unique identification number based on its physical location; (c) LD-blocks at the 15q25.1 locus visualized through HAPLOVIEW using LD estimates in the HapMAP-2 CEU reference population.

Figure 2. Several variants at chromosome 15q21.1 have stronger effects on CHD risk in "never-smokers" compared to "ever-smokers". Variants with the strongest interaction P-value are displayed.

Figure 3. Step-wise conditional analysis of genetic variation at the chromosome 15q21.1 locus with CHD (red triangles) and smoking behavior (cigarettes per day, CPD; grey circles). At the chromosome 15q21.1 locus, analyses adjusted for rs7178051 and rs11072794 completely attenuated the gene-CHD associations whereas gene-smoking remained unchanged. Analyses adjusted for rs1051730 and rs684513 completely attenuated the gene-smoking associations whereas gene-CHD effect remained unchanged.

Figure 4. Analyses mutually adjusted for rs7178051, rs11072794, rs1051730 and rs684513 at 15q21.1 on CHD and smoking behavior; gene-CHD interaction analyses were only found significant for rs7178051.

Figure 5. Genome browser view of regulatory features at rs7178051 on Chr15q21.1. ChIP-seq experiments were performed on confluent HCASMC for TCF21, Jun, JunD, CEBP and H3K4me1, H3K27me3, H3K27ac. DNAaseI hypersensitivity data for human AoSMC were acquired from the ENCODE project. Human aortic tissue H3K4me1, H3K9me3, H3K27me3, and H3K36me3 ChIP-seq data were acquired from the NIH Roadmap Epigenomics Project. HCASMC = human coronary artery smooth muscle cells; AoSMC = human aortic smooth muscle cells.

Figure 6. (a) *ADAMTS7* and *CHRNB4-A3-A5* mRNA levels were measured in HCASMC. Cells were cultured to confluence, total RNA was extracted and cDNA generated. q-PCR was performed for *ACTB, GAPDH, TBP, ADAMTS7, CHRNB4, CHRNA3, CHRNA5* (95°C 15s, 60°C 1min). Delta Cts

were calculated as follows: $(Ct_{ACTB} + Ct_{GAPDH} + Ct_{TBP})/3 - Ct_{TARGET\ GENE}$). Fold changes are derived from delta delta Cts based on formula $FC = 2^{-dCt}$. (b) Confluent HCASMC were exposed to cigarette smoke extract. Serum starved (x24 hrs.) confluent HCASMC were treated with 0.5% or 1.0% cigarette smoke extract (v/v) for 4, 12, and 24 hrs. in serum reduced conditions (0.5% FBS in DMEM). Total RNA was extracted, cDNA generated preparation and q-PCR performed for *ADAMTS7* by Taqman and normalized to *GAPDH*. The Average Ct for ADAMTS7 at baseline was 28.25. Results were presented as means ± SEM, and data were analyzed using Student's t-Test. (c) expression and eQTL Data from the GTEx consortium, the HapMap consortium (restricted to European populations), the Multiple Tissue Human Expression Resource (MuTHER) and in 147 donor HAoEC lines. Association of the independent lead variants identified in our conditional analyses with expression of *ADAMTS7* and genes in the *CHRNB4-A3-A5* cluster. A P-value threshold of 0.002 was set to account for multiple testing involved in the eQTL analyses.

Table-1. Novel genotype-smoking interaction findings in coronary heart disease at the chromosome 15q25.1 locus

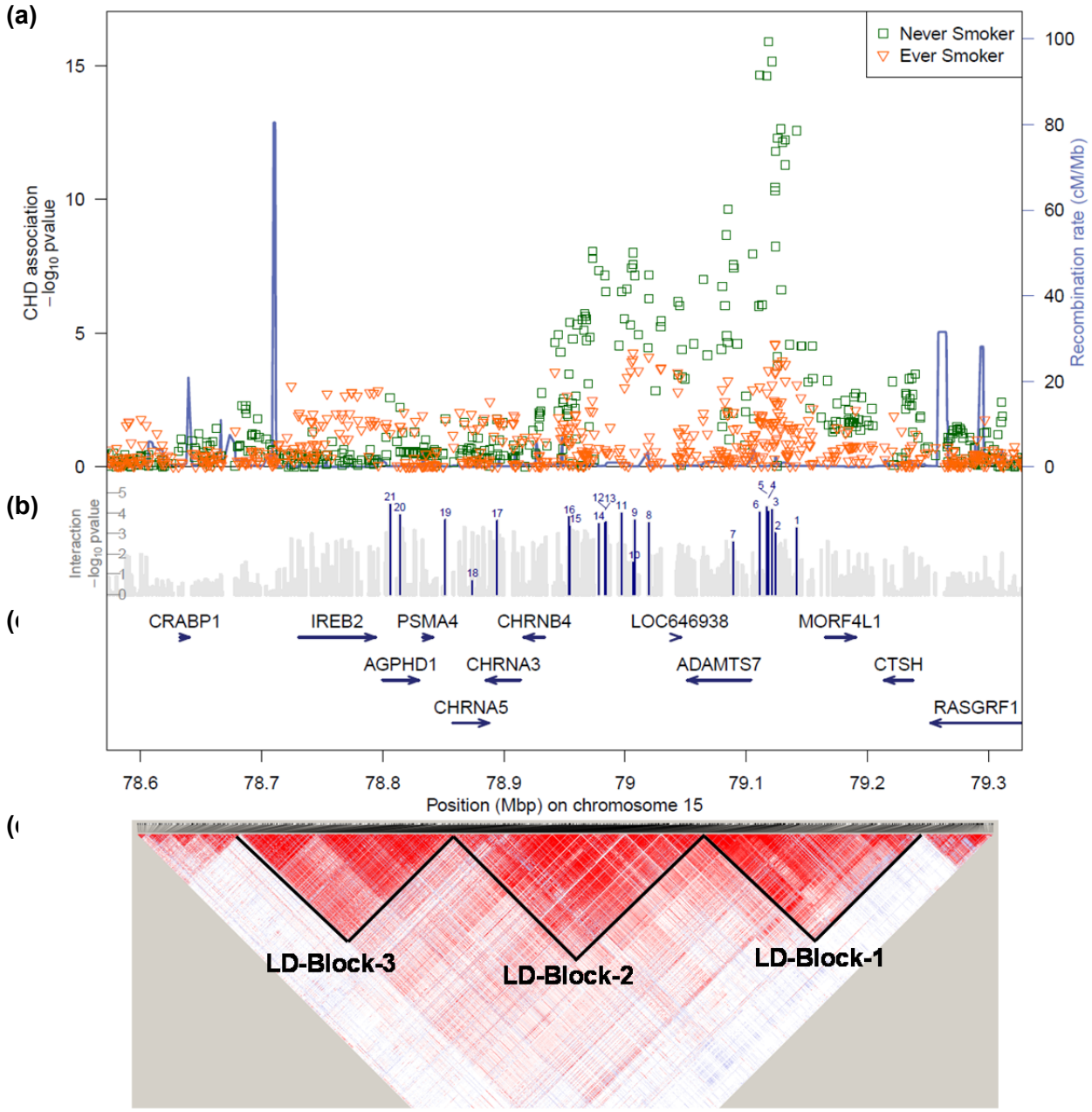
						Never Smokers						Ever Smokers					
Variant	**Association**	**allele**	**LD with** rs7178051*	**LD with** rs1051730^		**N cases**	**N controls**	**N Total**	**Beta (SE)**	**P-value**		**N cases**	**N controls**	**N Total**	**Beta (SE)**	**P-value**	**P-value interaction**
*rs7178051[4]	CHD (NPR)	T/C	-	0.22		21232	38713	59945	-0.13 (0.01)	1.30E-16		39585	40749	80334	-0.05 (.01)	2.49E-04	8.57E-05
^rs1051730[16]	SB (known)	A/G	0.22	-		20559	38198	58757	-0.04 (0.02)	0.02		38923	40371	79294	0.03 (0.01)	0.02	2.37E-04
Other variants on chr.15q25.1 with significant gene-smoking interactions on CHD																	
rs7173743[1]	CHD (Known)	C/T	0.61	0.18		21050	37955	59005	-0.11 (0.01)	2.73E-13		39044	39559	78603	-0.04 (0.01)	8.60E-04	9.29E-05
rs10083696[2]	CHD (Novel)	A/G	1.0	0.22		19721	36206	55927	-0.11 (0.02)	1.60E-12		38807	40018	78825	-0.05 (0.01)	2.72E-04	5.15E-05
rs7176187[3]	CHD (Novel)	T/C	1.0	0.24		21232	38713	59945	-0.12 (0.01)	7.02E-16		39585	40749	80334	-0.04 (0.01)	8.64E-04	6.93E-05
rs6495335[5]	CHD (Novel)	G/T	1.0	0.22		20144	37217	57361	-0.13 (0.02)	2.39E-15		36448	38203	74651	-0.04 (0.01)	1.69E-03	9.51E-04
rs4380028[6]	CHD (Known)	T/C	1	0.22		21232	38713	59945	-0.12 (0.01)	2.20E-15		39585	40749	80334	-0.04 (.01)	1.03E-03	5.44E-04
rs3825807[7]	CHD (Known)	G/A	0.52	0.43		17137	28633	45771	-0.09 (0.02)	2.82E-08		30071	29014	59086	-0.03 (0.01)	0.04	2.6E-03
rs3813565[8]	CHD (NPR)	T/G	0.43	0.56		19466	35830	55296	-0.08 (0.02)	5.08E-07		36642	37759	74401	-0.01 (0.01)	0.42	3.05E-04
rs11638490[9]	CHD (NPR)	T/C	0.44	0.51		20465	37897	58362	-0.08 (0.01)	6.90E-08		38533	39690	78223	-0.01 (0.01)	0.28	2.25E-04
rs11072791[11]	CHD (NPR)	A/C	0.44	0.51		19289	35944	55233	-0.08 (0.02)	2.83E-07		35245	36635	71880	-.005 (0.01)	0.68	1.06E-04
rs922692[12]	CHD (NPR)	A/C	0.44	0.50		20559	38198	58757	-0.08 (0.01)	2.81E-07		38923	40371	79294	-0.01 (0.01)	0.29	2.75E-04
rs11638372[13]	CHD (NPR)	T/C	0.44	0.50		21232	38713	59945	-0.08 (0.01)	6.92E-08		39585	40749	80334	-0.01 (0.01)	0.23	3.16E-04
rs4887077[14]	CHD (NPR)	T/C	0.44	0.50		21232	38713	59945	-0.08 (0.01)	4.71E-08		39585	40749	80334	-0.02 (0.01)	0.20	3.92E-05
rs12899135[15]	CHD (NPR)	G/A	0.39	0.56		20377	37440	57817	-0.07 (0.02)	3.97E-06		38382	39181	77563	-0.01 (0.01)	0.58	4.54E-04
rs684513[18]	SB (Known)	C/G	0.01	0.10		12517	21054	33572	-0.01 (0.02)	0.67		24641	24487	49129	0.03 (0.02)	0.18	0.08
rs2036527[19]	SB (Known)	A/G	0.17	0.90		20559	38198	58757	-0.04 (0.02)	0.02		38923	40371	79294	0.03 (0.01)	0.02	2.14E-04
rs10519203[20]	CHD (NPR)	G/A	0.19	0.93		21232	38713	59945	-0.04 (0.01)	5.93E-03		39585	40749	80334	0.03 (0.01)	0.03	1.27E-04
rs8034191[21]	SB (Known)	C/T	0.19	1.0		19251	32131	51382	-0.05 (0.02)	2.62E-03		34925	34047	68972	0.02 (0.01)	0.06	3.91E-05

CHD = coronary heart disease; SB = smoking behavior; NPR: Not a previously reported variant with disease risk

*lead variant in association with CHD in our dataset; ^ lead variant in association with SB

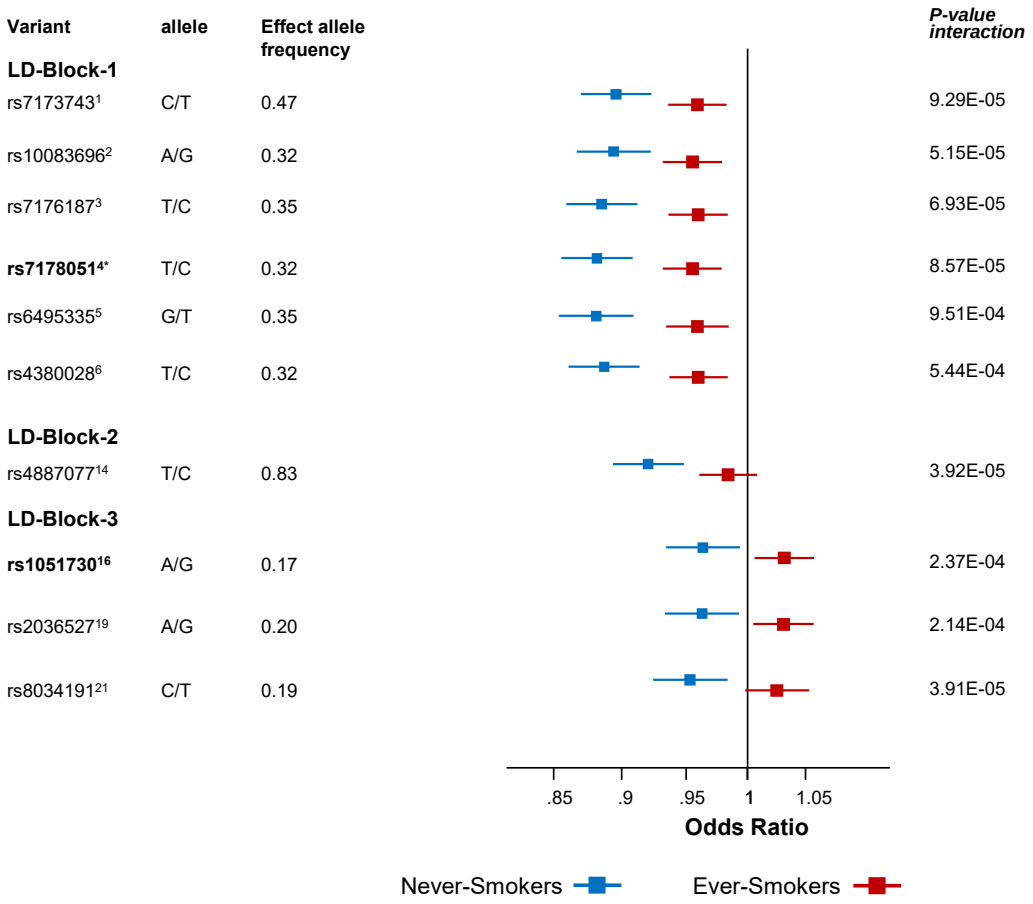
[1-21]each number refers to the physical location of the variant in figure-

Figure-1. Analyses of the chromosome 15q25.1 locus association with CHD stratified by smoking status and gene-smoking CHD interaction analyses



1-rs7173743; 2-rs10083696; 3-rs7176187; 4-rs7178051; 5-rs6495335; 6-rs4380028; 7-rs3825807; 8-rs3813565; 9-rs11638490; 10-rs11072794; 11-rs11072791; 12-rs922692; 13-rs11638372; 14-rs4887077; 15-rs12899135; 16-rs17487514; 17-rs1051730; 18-rs637137; 19-rs2036527; 20-rs10519203; 21-rs8034191. LD 1-3 indicate three separate linkage disequilibrium blocks in European ancestry at the chromosome 15q25.1 locus.

Figure-2. Multiple variants at chromosome 15q21.1 have stronger effects on CHD risk in "never-smokers" compared to "ever-smokers"



rs7178051 is the lead variant identified in association with CHD in our study population; whereas rs1051730 is the lead variant previously identified in association with smoking behavior .
Variants are ordered based on their base pair position in **Figure-1**.

Figure 3. Conditional analysis of genetic variation at the chromosome 15q21.1 locus with coronary heart disease (CHD; red triangles) and smoking behavior (cigarettes per day, CPD; grey circles)

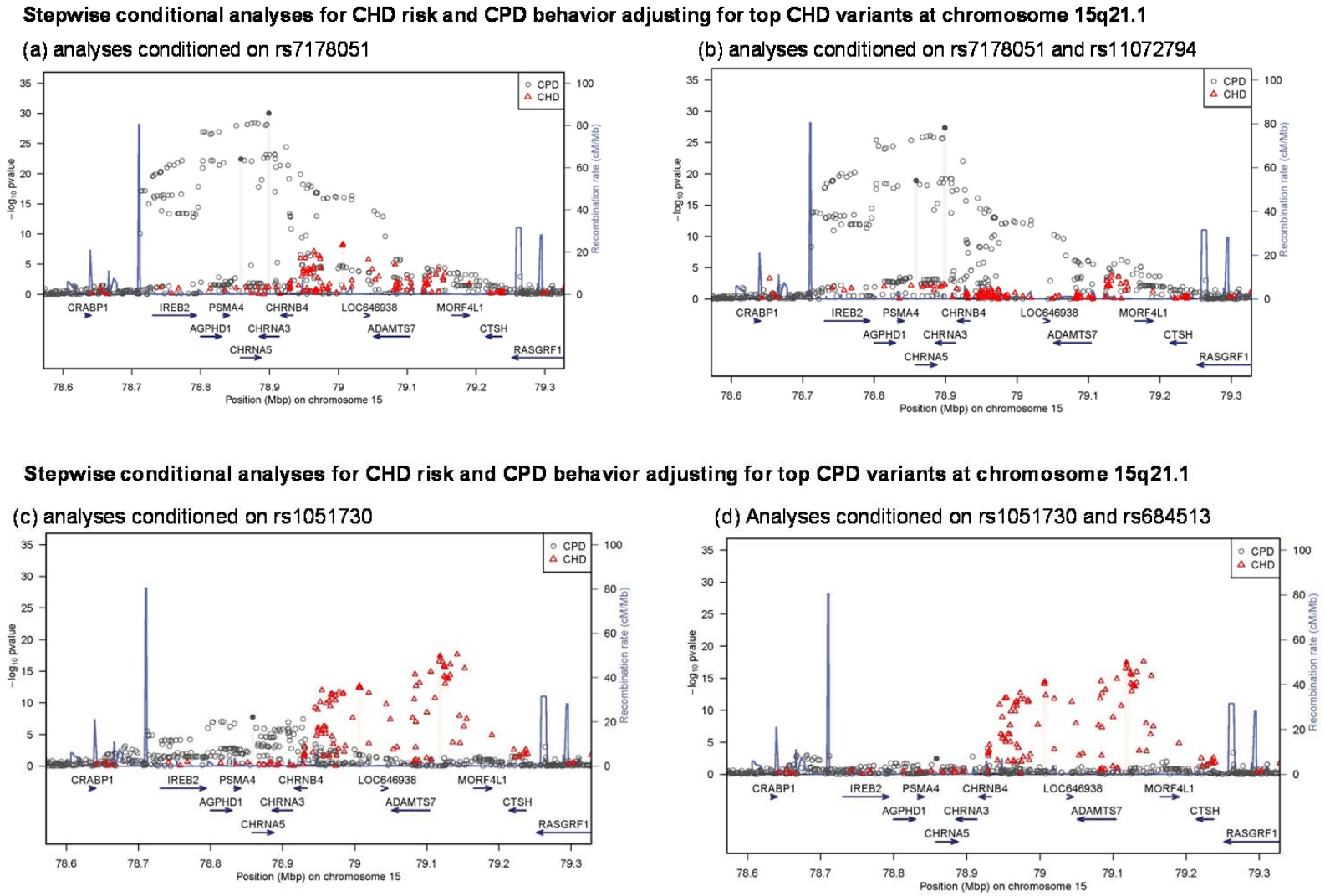
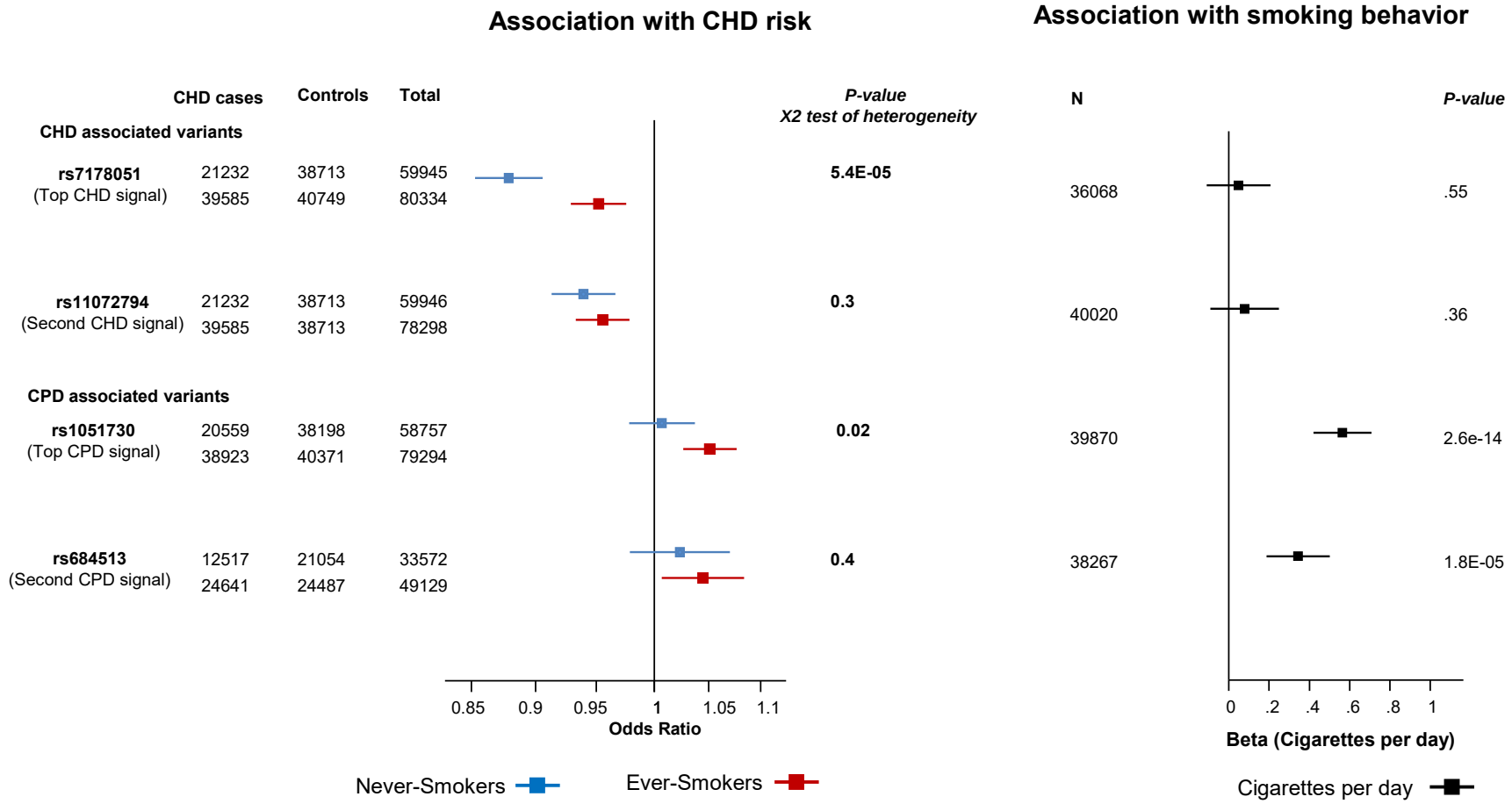


Figure 4. Mutually adjusted effects of 15q21.1 lead variants on coronary heart disease and smoking behavior



Gene-CHD and gene-smoking analyses for rs7178051 were adjusted for rs11072794, rs1051730 and rs684513;
Gene-CHD and gene-smoking analyses for rs11072794 were adjusted for rs7178051, rs1051730 and rs684513;
Gene-CHD and gene-smoking analyses for rs1051730 were adjusted for rs7178051, rs11072794 and rs684513;
Gene-CHD and gene-smoking analyses for rs684513 were adjusted for rs7178051, rs11072794 and rs1051730.

Figure-5. (a) Expression of *ADAMTS7* and *CHRNB4-A3-A5* mRNAs and (b) cigarette smoke extract (CSE) induction of *ADAMTS7* mRNA in primary human coronary artery smooth muscle cells

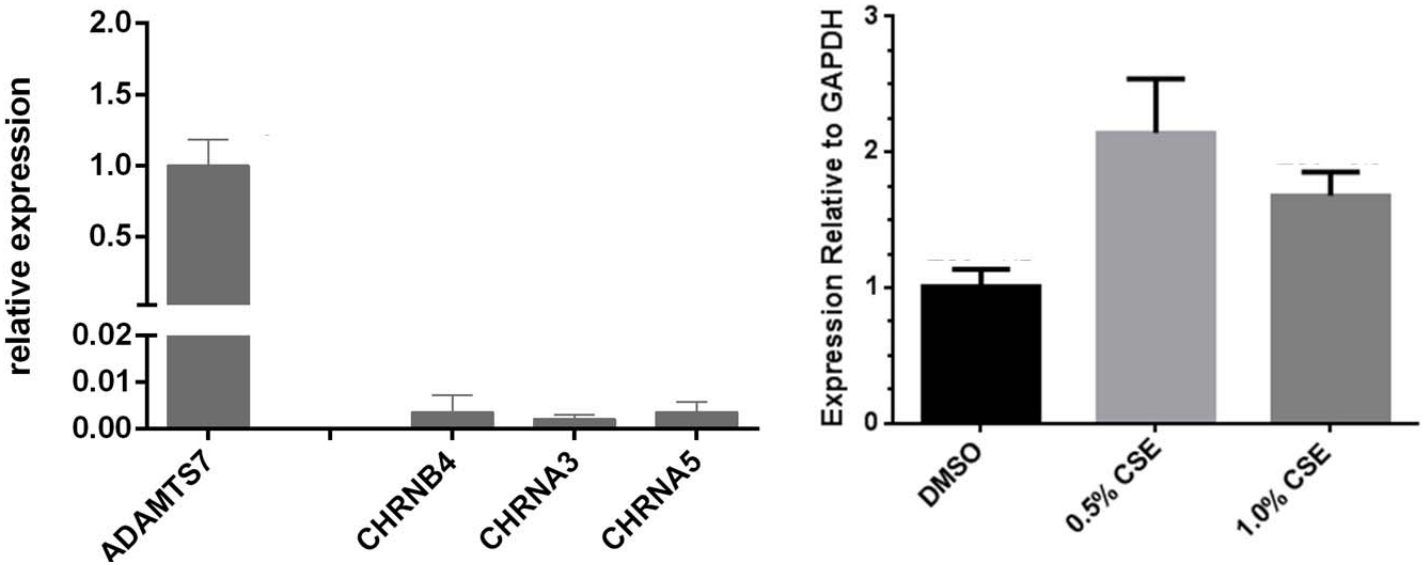


Figure-5c. Association of lead CHD and smoking behavior variants with candidate gene expression in available cells and tissues

			LCL in the MuTHER consortium (n=850)			**HAEC (n=147)**			**HapMap CEU LCL (n=109)**			**GTEx Skeletal muscle (n= 142)**			**GTEx Nerve Tibial (n=101)**	
Variant	Type	CHD direct.[#]	*ADAMTS7*	*CHRNA3-5*		*ADAMTS7*	*CHRNA3-5*		*ADAMTS7*	*CHRNA3-5*		*ADAMTS7*	*CHRNA3-5*		*ADAMTS7*	*CHRNA3-5*
rs7178051	Top CHD signal	-	**4.1e-4 (-)**	NS		**0.0029 (-)**	NS		NS	NS		NS	NS		NS	NS
rs11072794	Second CHD signal	-	**6.0E-21 (-)**	NS		NA	NS		**0.0013 (-)**	NS		NS	NS		NS	NS
rs1051730	Top CPD signal	-	NS	NS		NS	NS		NS	NS		NS	**6.9E-7 (-)**[1]		NS	**6.9E-7**[1]
rs684513	Second CPD signal	-	NS	NS		NS	NS		NS	NS		NS	**2.4E-7 (-)**[1]		NS	NS

[#]Direction of association for the effect allele on CHD; NS: Not significant (P-value < 0.002; Bonferroni correction for 20 tests); HAEC: Human Aortic Endothelial Cells; LCL: lymphoblastoid cell lines; 1 Association with *CHRNA5* expression;

Figure-6. Genome browser view of regulatory features at rs7178051 on Chr15q21.1